ADAPTIVE SEGMENTATION
OF TELEVISION IMAGES

5   This invention relates to video processing and more specifically to an adaptive segmentation system based upon characteristics such as color and texture, and in particular to sky detection.

      Segmentation of video images, such as television images, is a process wherein each frame of a sequence of images is subdivided into regions or segments. Each segment includes a cluster of pixels that encompass a region of the image with common properties or
10   characteristics. For example, a segment may be distinguished by a common color, texture, shape, amplitude range or temporal variation. Several methods are known for image segmentation using a process wherein a binary decision determines how the pixels will be segmented. According to such a process, all pixels in a region either satisfy a common criteria for a segment and are therefore included in the segment, or do not satisfy the criteria
15   and are completely excluded. While segmentation methods such as theses are satisfactory for some purposes, they are unacceptable for many others.

      In conventional methods of segmentation for grass detection, for example, the method based upon a probability distribution function for an expected grass color and luminosity in the YUV domain is representative of a compromise between computational simplicity and
20   algorithmic effectiveness. However, the three-dimensional Gaussian probability function defining the range of expected grass colors and luminosities in the YUV domain had increased expected values which were broad enough to account for possible variations of grass colors from scene to scene. This has the undesired side effect of increasing the false detection rate and declaring non-grass areas "grass." The same false detection problem arises
25   when the probability function is applied to methods for detecting other similar areas,. such as sky areas. In addition, bodies of water may at times be classified as sky areas, for example.

      Hence, there is a need for a method and system for adaptively segmenting video images that reduces the false classification of areas within the video images.

      A method and system for adaptively segmenting pixel elements in an image frame is
30   disclosed. The method comprises the steps of segmenting pixel elements into a at least one first region based on a selection criteria, refining the selection criteria based on information associated with each of the pixel elements within an associated first region and segmenting

the image pixel elements into at least one second region based on said refined selection criteria.

In the drawings:

Figure 1 illustrates a block diagram of an exemplary adaptive segmentation process in accordance with the principles of the present invention;

Figure 2 illustrates a block diagram of an exemplary process for determining an initial segmentation probability function;

Figure 3 illustrates a flow chart of an exemplary process for determining an updated probability function in accordance with the principles of the invention;

Figure 4 illustrates a flow chart of an exemplary process for determining an updated color probability function in accordance with the principles of the invention;

Figure 5 illustrates a flow chart of an exemplary process for determining pixels used in obtaining updated probability functions in accordance with the principles of the invention; and

Figure 6 illustrates a illustrates a system for executing the processing depicted in Figures 1-5.

It is to be understood that these drawings are solely for purposes of illustrating the concepts of the invention and are not intended as a definition of the limits of the invention. The embodiments shown in Figures 1 through 6 and described in the accompanying detailed description are to be used as illustrative embodiments and should not be construed as the only manner of practicing the invention. The same reference numerals, possibly supplemented with reference characters where appropriate, have been used to identify similar elements.

Video images may have significant areas or segments that may be identified as having substantially the same characteristics, e.g., color, luminese, texture. For example, a segment of an image may contain information related to a sky, i.e., blue color, smooth texture. Similarly, fields of grass may be identified by its green color and semi-smooth texture. Identification of areas, or segments of video images are more fully discussed in the commonly assigned, co-pending patent application serial no ____, entitled "Automatic Segmentation-based Grass Detection for Real-Time Video," and commonly assigned co-pending patent application serial no. ____, entitled, System and Method for Performing Segmentation-Based Enhancements of a Video Image."

Figure 1 illustrates a block diagram 100 of an exemplary adaptive segmentation process in accordance with the principles of the invention. In this embodiment, an initial

segmentation probability function is determined at block 110. As will be discussed, the initial probability function may be determined as a function of one or more probability functions including position, color and texture. At block 120, an updated position probability function is determined at block 120. At block 130, an updated color probability function is determined and at block 140 an updated texture probability function is determined. At block 150, an updated probability function is determined. The updated probability function is representative of a composite of the updated probability functions. At block 160, the image is re-evaluated using the updated probability function. In another aspect, the processing and the refinement of the probability distribution functions may be performed in parallel.

Figure 2 illustrates an exemplary process 110 for determining an initial probability function for segmentation. In this exemplary process, an initial position probability function is determined at block 210, an initial color probability function is determined at block 220 and an initial texture probability function is determined at block 230. At block 240, an initial segmentation probability function is determined in relation to the determined individual probability functions. With particular application to those areas of an image that may be related to the sky, a position function may assume that the sky is conventionally near a top of the image. Accordingly, a position probability function may be determined as:

$$P_{position} = e^{-(\frac{L}{\#lines})^2} \qquad\qquad [1]$$

where L=line number, starting from 0 at the top, and
#lines=the total number of scan lines per frame.

Similar probability distributions may be determined for other known regions, such as grass, water, faces, etc. In one aspect, the position probability distribution may be set to 1, i.e., uniform distribution, to indicate that no preference in position may be assumed or determined. In this case, the entire image may be associated with the known region.

An initial color probability distribution of the sky may be represented as:

$$P_{color} = e^{-\left[\left(\frac{y-y0}{\sigma_y}\right)^2 + \left(\frac{u-u0}{\sigma_u}\right)^2 + \left(\frac{v-v0}{\sigma_v}\right)^2\right]} \qquad\qquad [2]$$

where, initial starting values for sky detection may be set, on a
scale of 0-255, as:
$y0=210$, $\sigma y=130$;

$$u0=150, \sigma u = 40; \text{ and}$$

$$v0=100, \sigma v=40.$$

These parameters are determined empirically by examining a large number of sky images. However, it should be understood that other initial values may be used without altering the processing or the scope of the invention. Further, one skilled in the art will recognize that similar initial y, u, and v values for other image regions, such as grass, water, faces, etc, may be determined.

An initial texture probability function may be determined as:

$$P_{texture} = e^{-0.2*(t-t0)^2} \quad \text{for } t > t0 \text{ and} \qquad [3]$$

$$=1; \qquad \text{for } t < t0$$

where t0=10 for low noise; and

=40 for SNR=26dB; and

t is the sum of the absolute differences of 5 adjacent horizontal luminance values of a running window centered at the current pixel.

Similar probability distributions may be determined for other textures. In one aspect, the textual probability distribution may be set to 1, i.e., uniform distribution, to indicate that no preference in texture may be assumed or determined. In this case, the entire image may be associated with the known texture.

An initial probability function may be determined as:

$$P=P_{color}*P_{position}*P_{texture} \qquad [4]$$

Pixel elements matching or satisfying the selection criteria, as represented by P may be broadly classified, identified or associated with a known region of the image. In this manner, the broad and not very selective probability function reduces the chance of not detecting pixels within a desired region of interest.

Although, the probability function shown in equation 4 is determined in association with probability functions associated with color, position and texture, it will be understood by those skilled in the art that the probability function, P, may be similarly determined based

4

only on a single or any combination of the probability functions discussed or other characteristics of an image.

Figure 3 illustrates a flow chart of the exemplary process 120 for updating or refining the position probability function shown in Figure 1 in accordance with the principles of the invention. In this exemplary process, those pixels in the image satisfying a known position criteria are identified and tagged at block 300. Determination and identification of pixels satisfying a known threshold criteria associated with position will be discussed in more detail with regard to Figure 5.

At block 310, an initial scan line value is established or set. At block 320, a determination whether the number of pixels with a scan line satisfying the known position criteria is greater than a known threshold value. If the answer is in the negative, than a next scan line is obtained a block 350. A determination is then made at block 345 whether all the scan lines have been evaluated. If the answer is in the negative, then processing continues at block 320 to determine whether the number of pixels in the selected or obtained next scan line is greater than a known positional threshold.

Returning to block 320, if the answer affirmative, the scan line number that has a number of pixels satisfying the position criteria is saved or recorded for further processing at block 330. A next/subsequent scan line is selected or obtained at block 340 for processing.

Returning to block 345, if the answer is affirmative, i.e., all scan lines have been processed, the mean scan line value of the recorded or stored scan line values is determined at block 350. Using the mean scan line value the positional probability function is updated at block 360.

In one aspect of the invention, scan lines may be numbered from top-to bottom and the pixels in each scan line numbered left-to-right. In this manner each pixel may be uniquely identified. Furthermore, a next/subsequent scan line or pixel may be selected or obtained from a preceding scan line or pixel by incrementing a scan line or pixel number. Similar methods of identifying and selecting scan lines and associated pixel elements are well-known and need not be discussed in detail herein.

In one embodiment of the invention, a positional threshold criteria may be selected with regard to a probability function as:

$$K_l*\text{maximum}(P) \hspace{4cm} [5]$$

where K_1 is a known percentage of the maximum probability
function.

In a preferred embodiment, K_1 is set to 0.5. Hence, in the preferred embodiment, those
pixels (i) in each scan line (j) satisfying the criteria

$$P_{ij} > 0.5 * \text{maximum}(P) \qquad\qquad [6]$$

are identified, stored or retained for subsequent processing.

An updated positional probability function may be then determined as:

$$P_{\text{position2}} = \frac{(3*s - L)}{(3*s)} > 0 \qquad\qquad [7]$$

where L=scan line number (0 at the top); and
       s=is the mean scan line value determined in block 350.

In another aspect of the invention, the scan line values are stored when it is
determined that a sufficient percentage of pixels within a scan line satisfy the criteria shown
in Equation 5. For example, a scan line is saved when the number of pixels satisfying the
preferred criteria shown in Equation 6 exceeds for example, three percent (3%) of the total
number of pixels in a selected scan line. Accordingly, in this aspect of the invention, a scan
line is stored or recorded when:

       # pixels satisfying equation 6 > (Total #. pixel in scan line)/K_2 [8]

where K_2 is <= 32.

In a preferred embodiment K_2 is equal to 32.

Figure 4 illustrates a flow chart of an exemplary process 130 depicted in Figure 1 for
updating a color probability function. In this exemplary process a number of pixels in each
scan line satisfying a known color-related threshold value are determined at block 300.

At block 410, a mean value corresponding to the each color level associated with each
pixel satisfying the known color threshold is then determined. At block 420, an updated
color probability function may be determined using the determined mean color values.

Determination and identification of pixels satisfying a known threshold criteria associated with color will be discussed in more detail with regard to Figure 5.

In one embodiment of the invention, a color criteria may be selected with regard to a probability function as:

$$K_3*\text{maximum}(P) \qquad\qquad [9]$$

where K_3 is a known percentage of the maximum probability function.

In a preferred embodiment, K_3 is set to 0.95. Hence, in the preferred embodiment, those pixels (i) in each scan line (j) satisfying the criteria

$$P_{ij}>0.95*\text{maximum}(P) \qquad\qquad [10]$$

are identified, stored or retained for subsequent processing.

Mean color values associated with each pixel satisfying the color criteria of equation 10 may be determined as:

$$y_1=\frac{\sum_1^N y_{ij}}{N}, \ u_1=\frac{\sum_1^N u_{ij}}{N}, \ v_1=\frac{\sum_1^N v_{ij}}{N} \qquad\qquad [11]$$

where y_{ij}, u_{ij}, v_{ij} are representative of the color levels of the ij^{th} pixel; and

N is the total number of pixels satisfying the color criteria.

An updated color probability function may then be determined by:

$$P_{color2}=e^{-\left[\left(\frac{y-y1}{k*\sigma_y}\right)^2+\left(\frac{u-u1}{k*\sigma_u}\right)^2+\left(\frac{v-v1}{k*\sigma_v}\right)^2\right]} \qquad\qquad [12]$$

It should be appreciated by those skilled in the art, that the denominators of each term in the exponent have been multiplied by a factor k, wherein k is less than one (1). Use of the factor k is advantageous as it results in a smaller sigma value and consequentially to a distribution that is more peaked or concentrated. In this manner, the selection of pixels in a region is limited by the narrower or concentrated distribution function. In a preferred embodiment, k may be equal to 0.5.

7

An updated texture probability function, $P_{texture2}$, may then be determined as:

$$P_{texture2} = e^{-0.2*(tt)^2}$$ [13]

where tt is the absolute difference of luminance values of a current pixel and the following one on the same line.

Although, an updated texture probability is determined with regard to an difference in luminance values as described in equation 13, it should be understood that an updated texture probability may determined using pixels satisfying a known texture-related threshold value similar to that disclosed with regard to an updated position or color probability density, as discussed previously.

An updated probability function may then be determined as:

$$P_u = P_{color2}*P_{position2}*P_{texture2}$$ [14]

The updated probability function P_u may then be used to re-classify each pixel in the image to refine the determination of those pixels within desired or designated areas or regions of interest. For example, updated probability distribution function P_u may be used to refine the determination of those pixels in, for example, sky, grass, or face regions of an image.

Figure 5 illustrates a flow chart 500 of the exemplary process 300, shown in Figure 3, for determining pixel elements that satisfy a threshold criteria associated with a positional probability function, and in Figure 4 for determining pixel elements that satisfy a threshold criteria associated with a color probability function. Hence, when an updated positional probability function is to be determined, a threshold criteria may be determined in accordance with equation 5, and in addition equation 8. And, when an updated color probability function is to be determined, then the threshold criteria may be determined in accordance with equation 9.

In flow chart 500, an initial scan line value is set or an initial scan line is selected at block 510. Preferably, an initial scan line is set to the top-most line, i.e., zero line, of the image. At block 520, an initial pixel position within the selected scan line is selected. At block 530 a determination is made whether the probability associated with the selected pixel is greater than a known threshold value or criteria. If the answer is in the affirmative, the identification of the pixel satisfying the threshold criteria is stored or recorded at block 540. However, if the answer is negative, then a next/subsequent pixel in the selected scan line is selected at block 550. At block 560, a determination is made whether all pixels on the selected scan line have been processed. If the answer is in the negative, then processing

continues at block 530 to determine whether the next/subsequent pixel selected is greater than the known threshold.

However, if the answer at block 560 is affirmative, then a next/subsequent scan line is selected at block 570. At block 580, a determination is made whether all the scan lines in the image have been processed. If the answer is in the negative, then processing continues at block 520 to select a pixel element associated with the selected next/subsequent scan line. If, however, the answer to the determination at block 580 is in the affirmative, then process is completed.

Figure 6 illustrates an exemplary embodiment of a system 600 that may be used for implementing the principles of the present invention. System 600 may represent a real-time receiving system, such as an SDTV or HDTV television, a desktop, laptop or palmtop computer, a personal digital assistant (PDA), a video/image storage apparatus such as a video cassette recorder (VCR), a digital video recorder (DVR), a TiVO apparatus, etc., as well as portions or combinations of these and other devices. System 600 may contain one or more input/output devices 602, processors 603 and memories 604. I/O devices may access or receive information from one or more sources 601 that contain video images. Sources 601 may be stored in permanent or semi-permanent media such as a television receiving system, a VCR, RAM, ROM, hard disk drive, optical disk drive or other video image storage devices. Sources 601 may alternatively be accessed over one or more network connections 625 for receiving video from a server or servers over, for example a global computer communications network such as the Internet, a wide area network, a metropolitan area network, a local area network, a terrestrial broadcast system (Radio, TV), a cable network, a satellite network, a wireless network, or a telephone network, as well as portions or combinations of these and other types of networks.

Input/output devices 602, processors 603 and memories 604 may communicate over a communication medium 606. Communication medium 606 may represent, for example, a bus, a communication network, one or more internal connections of a circuit, circuit card or other apparatus, as well as portions and combinations of these and other communication media. Input data from the sources 601 is processed in accordance with one or more programs that may be stored in memories 604 and executed by processors 603. Processors 603 may be any means, such as general purpose or special purpose computing system, or may be a hardware configuration, such as a laptop computer, desktop computer, handheld computer, dedicated logic circuit, integrated circuit. Processors 603 may also be

Programmable Array Logic (PAL), Application Specific Integrated Circuit (ASIC), etc., which may be hardware "programmed" to include software instructions that provide a known output in response to known inputs.

In one embodiment, the coding employing the principles of the present invention may
5   be implemented by computer readable code executed by processor 603. The code may be stored in the memory 604 or read/downloaded from a memory medium such as a CD-ROM or floppy disk (not shown). In a preferred embodiment hardware circuitry may be used in place of, or in combination with, software instructions to implement the invention. For example, the elements illustrated herein may also be implemented as discrete hardware
10   elements that are operable to perform the operations shown using coded logical operations or by executing hardware executable code.

Data from the source 601 received by I/O device 602 after processing in accordance with one or more software programs operable to perform the functions illustrated in Figures 2 and 3, which may be stored in memory 604 and executed by processor 603 may then be
15   transmitted over network 630 to one or more output devices represented as TV monitor 640, storage device 645 or display 650. As will be appreciated, TV monitor 640 may be an analog or digital TV monitor.

The term computer or computer system may represent one or more processing units in communication with one or more memory units and other devices, e.g., peripherals,
20   connected electronically to and communicating with the at least one processing unit. Furthermore, the devices may be electronically connected to the one or more processing units via internal busses, e.g., ISA bus, microchannel bus, PCI bus, PCMCIA bus, etc., or one or more internal connections of a circuit, circuit card or other device, as well as portions and combinations of these and other communication media or an external network, e.g., the
25   Internet and Intranet.